Two Remarks on Matrix Exponentials

Edgar M. E. Wermuth
Zentralinstitut für Angewandte Mathematik (ZAM)
Kernforschungsanlage Jülich GmbH (KFA)
Postfach 1913
D-5170 Jülich, West Germany

Submitted by Ludwig Elsner

ABSTRACT

It is shown that for two square matrices A and B with algebraic elements, $e^A e^B = e^B e^A$ if and only if $AB = BA$. Moreover, the in some sense best possible perturbation inequality

$$\|e^A - e^B\|_2 \leqslant \|e^A\|_2 (e^{\|A-B\|_2} - 1)$$

for a normal matrix A is proven.

---

1. NONCOMMUTING AND COMMUTING EXPONENTIALS

It is a plain and well-known fact that if $AB = BA$ holds for two square matrices A and B, we also have $e^A e^B = e^B e^A = e^{A+B}$.

In this note we want to prove a partial converse; we shall show that under certain circumstances $e^A e^B = e^B e^A$ implies $AB = BA$. Of course, no general conclusion of this type is possible, as can be seen by simple examples: If

$$A_1 = \begin{pmatrix} \pi i & 0 \\ 0 & -\pi i \end{pmatrix}, \qquad B_1 = \begin{pmatrix} 0 & 1 \\ 0 & -2\pi i \end{pmatrix}, \qquad B_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix},$$

we have

$$A_1 B_1 \neq B_1 A_1, \qquad \text{but} \quad e^{A_1} e^{B_1} = e^{B_1} e^{A_1} = e^{A_1 + B_1},$$

655 Avenue of the Americas, New York, NY 10010 0024-3795/89/$3.50

and

$$e^{A_1}e^{B_2} = e^{B_2}e^{A_1} \neq e^{A_1+B_2}.$$

If now

$$A_2 = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} \quad \text{with} \quad a \neq b,$$

we get

$$e^{A_2}e^{B_2} = \begin{pmatrix} e^a & e^a \\ 0 & e^b \end{pmatrix}, \quad e^{B_2}e^{A_2} = \begin{pmatrix} e^a & e^b \\ 0 & e^b \end{pmatrix}, \quad e^{A_2+B_2} = \begin{pmatrix} e^a & \dfrac{e^a - e^b}{a-b} \\ 0 & e^b \end{pmatrix}.$$

Since $e^z - z$ assumes every complex value infinitely often (Picard's theorem and the periodicity of e^z), we see, taking $b - a = z$, that for a lot of values a and b

$$e^{B_2}e^{A_2} \neq e^{A_2}e^{B_2} = e^{A_2+B_2}.$$

On the other hand, for all but a countable number of values $a - b$ there holds no equality between any two of the three expressions $e^{B_2}e^{A_2}$, $e^{A_2}e^{B_2}$, $e^{A_2+B_2}$.

In *SIAM Rev.* D. S. Bernstein (1988) posed a problem that refers to an article of M. Fréchet's (1952). Fréchet assumed (see p. 12 and p. 18 of his paper) that $e^Ae^B = e^{A+B}$ always entails $e^Ae^B = e^Be^A$. This is seen to be incorrect by one of the examples given here. Restriction to real matrices does not change the situation, since our complex two by two examples can be read as real four by four examples using the identification of a complex number $x + iy$ with the real matrix

$$\begin{pmatrix} x & -y \\ y & x \end{pmatrix}.$$

Bernstein's problem of proving or disproving Fréchet's assumption for real matrices is thus solved negatively. An explicit example: If $z = a + ib$ is a solution of $e^z - z = 1$, e.g. $a = 2.088843\ldots$ and $b = 7.461489\ldots$, then for

$$A = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & a & -b \\ 0 & 0 & b & a \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

we get

$$e^A e^B = e^{A+B} = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & a+1 & -b \\ 0 & 0 & b & a+1 \end{pmatrix}, \quad e^B e^A = \begin{pmatrix} 1 & 0 & a+1 & -b \\ 0 & 1 & b & a+1 \\ 0 & 0 & a+1 & -b \\ 0 & 0 & b & a+1 \end{pmatrix}.$$

It should be mentioned[1] that Fréchet corrected his error (which also was detected by W. Givens in a review of Fréchet's paper) in a subsequent note (1953), and several other authors published studies on the solutions of $e^x e^y = e^{x+y}$.

For noncommuting exponents thus almost everything being possible, we nevertheless have the following

Theorem 1. *If A and B are square matrices with algebraic elements, then $e^A e^B = e^B e^A$ if and only if $AB = BA$.*

Proof. Only one direction needs to be proven, the other one being obvious.

Let $m(\lambda) = \prod_{j=1}^{k}(\lambda - \lambda_j)^{\mu_j}$ be the minimum polynomial of A. Lindeman's theorem on the transcendence of π implies that no two zeros of m differ by an integral multiple of $2\pi i$. Hence, using Hermite's interpolation formula, we can choose a polynomial f such that for $g = f \circ \exp$ the relations $g(\lambda_j) = \lambda_j$ $(1 \leqslant j \leqslant k)$, $g'(\lambda_j) = 1$ for $\mu_j > 1$ $(1 \leqslant j \leqslant k)$, and $g^{(\nu)}(\lambda_j) = 0$ $(1 \leqslant j \leqslant k,\ 1 < \nu < \mu_j)$ hold. By well-known properties of matrix functions we conclude that $g(A) = A = f(e^A)$, whence A is representable as a polynomial in e^A. Since this is true for B and e^B, too, we see that in the case of $e^A e^B = e^B e^A$ the relation $AB = BA$ necessarily follows. ■

The theorem implies that every pair of noncommuting algebraic (in particular rational) matrices A and B violates the functional equation of the exponential, since at least one of the products $e^A e^B$ or $e^B e^A$ differs from e^{A+B}. As is obvious from the proof, the equivalence stated in the theorem also holds for nonalgebraic matrices A and B if, e.g.,

$$\lim_{n \to \infty} \sqrt[n]{\|A^n\|} < \pi \quad \text{and} \quad \lim_{n \to \infty} \sqrt[n]{\|B^n\|} < \pi.$$

[1]Communicated to the author by Professor M. S. Klamkin.

2. A PERTURBATION INEQUALITY FOR NORMAL EXPONENTIALS

In connection with the case of noncommuting matrices the classical Lie product formula

$$\lim_{n \to \infty} \left(e^{A/n} e^{B/n}\right)^n = e^{A+B}$$

deserves more popularity in linear algebra courses; it is strange that for a proof, simple and elementary as it is, we have to refer to quite advanced textbooks like Reed and Simon (1972) or Varadarajan (1984). We use it here to derive a best possible perturbation inequality for normal matrix exponentials which is slightly better than the estimate mentioned in van Loan (1977), Moler and van Loan (1978), and Golub and van Loan (1983). Since the argument is not restricted to the finite dimensional case, we have chosen a Hilbert space formulation. Thus in the matrix case the norm has to be understood as being Euclidean or unitary.

THEOREM 2. *If A is a normal bounded linear operator on a Hilbert space and B is another bounded linear operator on that space, then*

$$\|e^A - e^B\| \leqslant \|e^A\|\left(e^{\|A-B\|} - 1\right).$$

Proof. Observing

$$\begin{aligned}\frac{d}{dt}\left(e^{A(1-t)} e^{Bt}\right) &= -Ae^{A(1-t)} e^{Bt} + e^{A(1-t)} B e^{Bt} \\ &= e^{A(1-t)}(B - A) e^{Bt},\end{aligned}$$

we conclude

$$e^A - e^B = \int_0^1 e^{A(1-t)}(A - B) e^{Bt}\, dt.$$

Hence we get the estimate

$$\|e^A - e^B\| \leqslant \|A - B\| \int_0^1 \|e^{A(1-t)}\| \, \|e^{Bt}\| \, dt.$$

Since for normal operators A we have $\|A^n\| = \|A\|^n$, by a straightforward continuity argument we see $\|e^{At}\| = \|e^A\|^t$ for every $t \geqslant 0$.

Now we use the Lie product formula to get

$$\begin{aligned}\|e^{Bt}\| &= \lim_{n\to\infty} \left\| \left(e^{(B-A)t/n} e^{At/n} \right)^n \right\| \\ &\leqslant \|e^A\|^t \limsup_{n\to\infty} \|e^{(B-A)t/n}\|^n \\ &\leqslant \|e^A\|^t e^{\|A-B\|t},\end{aligned}$$

and hence by the former estimate we have

$$\|e^A - e^B\| \leqslant \|A - B\| \, \|e^A\| \int_0^1 e^{\|A-B\|t}\, dt = \|e^A\|(e^{\|A-B\|} - 1). \qquad \blacksquare$$

If A and B are arbitrary *commuting* operators on a Banach space, then the same estimate trivially follows from

$$e^A - e^B = e^A \int_0^1 (A - B) e^{(B-A)t}\, dt.$$

The inequality is best possible in the sense that a relation $\|e^A - e^B\| / \|e^A\| \leqslant f(\|A - B\|)$ holding for a given normal operator A and an arbitrary B implies $f(x) \geqslant e^x - 1$ $(x \geqslant 0)$; in the abovementioned literature $f(x) = xe^x$ is used instead.

ADDENDUM

The papers by Morinaga and Nôno (pointed out to the author by Professor J. L. Brenner after completion of the present paper) contain a very thorough discussion of the solutions of $e^x e^y = e^{x+y}$, especially of those of orders two and three. Among other things, the authors prove that there are no real order two counterexamples to Fréchet's false claim. The 1950 paper also contains a result similar to the one mentioned here at the end of Section 1. However, the way of derivation is more complicated than a proof based on that of Theorem 1 in the present note.

For valuable comments the author thanks his colleagues J.-Fr. Hake, W. Homberg, W. Nagel, and P. Weidner, as well as M. Geck (TH Aachen).

The author is indebted to Professor M. S. Klamkin for communicating a list of papers (on the solutions of $e^x e^y = e^{x+y}$) due to Professor F. Leite (Coimbra) and to Professor J. Brenner (Palo Alto).

Moreover, he would like to thank the referee for some helpful suggestions.

REFERENCES

Bernstein, D. S. 1988. Problem 88-1, *SIAM Rev.* 30:123.

Fréchet, M. 1952. Les solutions non commutables de l'équation matricielle $e^X e^Y = e^{X+Y}$, *Rend. Circ. Mat. Palermo (2)* 1:11–27.

Fréchet, M. 1953. Rectification, *Rend. Circ. Mat. Palermo (2)* 2:71–72.

Golub, G. H. and van Loan, C. F. 1983, *Matrix Computations*, Johns Hopkins U.P., Baltimore.

Huff, C. W. 1953. On pairs of matrices (of order two) A, B satisfying the condition $e^A e^B = e^{A+B} \neq e^B e^A$, *Rend. Circ. Mat. Palermo (2)* 2:326–330.

Moler, C. and van Loan, C. F. 1978. Nineteen dubious ways to compute the exponential of a matrix, *SIAM Rev.* 20:801–836.

Morinaga, K. and Nôno, T. 1950. On the logarithmic functions of matrices I, *J. Sci. Hiroshima Univ. (A)* 14:107–114.

Morinaga, K. and Nôno, T. 1954. On the non-commutative solutions of the exponential equation $e^x e^y = e^{x+y}$, *J. Sci. Hiroshima Univ. (A)* 17:345–358.

Morinaga, K. and Nôno, T. 1954. On the non-commutative solutions of the exponential equation $e^x e^y = e^{x+y}$ II, *J. Sci. Hiroshima Univ. (A)* 18:137–178.

Reed, M. and Simon, B. 1972. *Methods of Modern Mathematical Physics I — Functional Analysis*, Academic, New York.

van Loan, C. F. 1977. The sensitivity of the matrix exponential. *SIAM J. Numer. Anal.* 14:971–981.

Varadarajan, V. S. 1984, *Lie Groups, Lie Algebra, and Their Representations*, Springer-Verlag, New York.

Received 6 June 1988; final manuscript accepted 1 September 1988